

The Square Variation of Rearranged Fourier Series

Allison Lewko

Mark Lewko*

Abstract

We prove that there exists a rearrangement of the first N elements of the trigonometric system such that the L^2 -norm of the square variation operator is at most $O_\epsilon(\log^{9/22+\epsilon}(N))$. This is an improvement over $O(\log^{1/2}(N))$ from the canonical ordering.

1 Introduction

Let $\Phi := \{\phi_n\}_{n=1}^N$ denote an orthonormal system (ONS) of functions from a probability space, \mathbb{T} , to \mathbb{R} . One is often interested, usually motivated by questions regarding almost everywhere convergence, in the behavior of the maximal partial sum operator

$$\mathcal{M}f(x) := \max_{\ell \leq N} \left| \sum_{n=1}^{\ell} a_n \phi_n(x) \right|.$$

For an arbitrary ONS, the Rademacher-Menshov theorem states that $\|\mathcal{M}f\|_{L^2} \ll \log(N)\|f\|_{L^2}$, where the $\log(N)$ factor is known to be sharp. However, one can do much better for many classical systems, for instance the well-known Carleson-Hunt inequality allows one to replace $\log(N)$ with an absolute constant in the case of the trigonometric system. More recently, there has been interest in variational refinements of these maximal results. We define the r -variation operator

$$\mathcal{V}^r f(x) := \left(\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{n \in I} a_n \phi_n(x) \right|^r \right)^{1/r}$$

where \mathcal{P}_N denotes the set of partitions of $[N]$ into subintervals. Clearly, $\mathcal{V}^r f$ is pointwise non-decreasing as r decreases, and $\mathcal{M}f \leq \mathcal{V}^r f$ for all $r < \infty$. While estimates involving the maximal operator typically imply statements regarding almost everywhere convergence, estimates involving the larger variational operators typically imply quantitative statements about the rate of convergence. In [10], it was shown that the Rademacher-Menshov theorem can be strengthened to $\|\mathcal{V}^2 f\|_{L^2} \ll \log(N)\|f\|_{L^2}$.

In the case of the trigonometric system, it was shown by Oberlin, Seeger, Tao, Thiele, and Wright [13] that $\|\mathcal{V}^r f\|_2 \ll_r \|f\|_2$ for $r > 2$ (strengthening the Carleson-Hunt inequality). In the case $r = 2$, one can deduce the inequality $\|\mathcal{V}^2 f\|_2 \ll \sqrt{\log(N)}\|f\|_2$ from the Carleson-Hunt inequality (see [10], Theorem 3). Moreover, the $\sqrt{\log(N)}$ can be shown to be sharp (see [13], section 2).

When considering questions regarding partial sums of an ONS, the ordering of the system plays a crucial role. For instance, Olevskii [14] has shown that any infinite complete ONS can be reordered in a manner such that almost everywhere convergence fails for some L^2 function. The related question of whether an ONS can be reordered in a manner such that almost everywhere

*M. Lewko is supported by a NSF Postdoctoral Fellowship, DMS-1204206.

convergence holds for every L^2 function is known as Kolmogorov's rearrangement problem and is one of the central open problems regarding orthonormal systems. Going further, Garsia has conjectured [6] (see also [5]) that given an ONS $\Phi := \{\phi_n\}_{n=1}^N$, one may find a permutation $\sigma(n) : [N] \rightarrow [N]$ such that the reordered system $\{\phi_{\sigma(n)}\}_{n=1}^N$ satisfies $\|\mathcal{M}f\|_2 \ll \|f\|_2$ where the implied constant is absolute and independent of the system. This is known to imply an affirmative solution to Kolmogorov's problem. As partial progress towards Garsia's conjecture, Bourgain [4] has shown (for uniformly bounded systems) that one may always find a permutation of $[N]$ such that $\|\mathcal{M}f\|_2 \ll \log \log(N) \|f\|_2$. In [10], this was strengthened to $\|\mathcal{V}^r f\|_2 \ll_r \log \log(N) \|f\|_2$ for $r > 2$. While these estimates have strong consequences for very general orthonormal systems, their conclusions are weaker than what is known to be true for most classical systems (such as the trigonometric system) in their canonical orderings.

With this in mind, it was asked in [10] if one could improve the inequality $\|\mathcal{V}^2 f\|_2 \ll \sqrt{\log(N)} \|f\|_2$ by reordering the first N elements of the trigonometric system. Here we provide an affirmative answer to this question by proving the following theorem:

Theorem 1. *Let $\Phi := \{\phi_n\}_{n=1}^N$ denote a ONS¹ uniformly bounded by some constant A . Let $\epsilon > 0, \gamma > 1$. Then, with probability at least $1 - cN^{-\gamma}$ (for some universal c), for a uniformly random permutation $\sigma : [N] \rightarrow [N]$, the system $\{\phi_{\sigma(n)}\}_{n=1}^N$ will satisfy*

$$\|\mathcal{V}^2 f\|_2 \ll_{A,\epsilon,\gamma} \log^{\frac{9}{22}+\epsilon}(N) \|f\|_2.$$

It turns out that treating the \mathcal{V}^2 operator requires a considerably more delicate analysis than the maximal and r -variation ($r > 2$) cases previously studied. Indeed, the Dudley-type chaining/covering number methods used in [4] and [10] alone seem incapable of achieving anything better than $\|\mathcal{V}^2 f\|_{L^2} \ll \log^{1/2}(N) \log \log(N) \|f\|_{L^2}$ (see [10]). The limitations of these methods have been previously overcome in the context of related problems, most notably Bourgain's solution to the $\Lambda(p)$ -problem [4], as well as Talagrand's alternate approach and generalizations [16]. The probabilistic component of our current work will use tools from both Bourgain and Talagrand's works, however additional complications enter as we will need to work with random subsets of a much greater density and derive stronger concentration bounds (see Section 7 for further discussion of these issues and an overview of this part of our argument). A careful combinatorial organization is also needed to reduce estimates for the \mathcal{V}^2 operator to questions amenable to these probabilistic methods. Here we rely, in part, on ideas from Taylor's work [19] on the path variation of Brownian motion.

We do not expect that the exponent $\frac{9}{22}$ is sharp. In the maximal case, it is known that Bourgain's estimate $\|\mathcal{M}f\|_2 \ll \log \log(N) \|f\|_2$ is the best one can achieve with a purely probabilistic argument (see remark 2 in [4]). It is consistent with our knowledge that probabilistic arguments might be able to achieve $\|\mathcal{V}^2 f\|_2 \ll \log \log(N) \|f\|_2$, although this would surely require a much deeper analysis. In [10] (see Theorem 6) it was shown that for any bounded ONS one may find a function f such that $\|\mathcal{V}^2 f\|_2 \gg \sqrt{\log \log(N)} \|f\|_2$, which gives a lower bound for any ordering. This fact is closely related to the law of the iterated logarithm (see [9]).

2 Preliminaries

We use the notation $x \ll y$, for instance, to mean that there exists an absolute constant C such that $x \leq Cy$. We similarly employ notation like $x \ll_p y$ to mean that there exists a constant C_p depending only on p such that $x \leq C_p y$. We will use $[N]$ to denote the set of the first N natural numbers $\{1, 2, \dots, N\}$.

¹We have defined an ONS to be real-valued, however the result follows for the complex-valued trigonometric system by applying the result to the real and imaginary parts separately.

Remark 1. Throughout this paper, we will consider orthonormal systems $\Phi := \{\phi_n\}_{n=1}^N$ uniformly bounded by a fixed constant A , meaning that $|\phi_n(x)| \leq A$ for all $n \in [N]$ and all $x \in \mathbb{T}$. We will refer to these simply as **bounded orthonormal systems**. Since A is fixed, we allow dependence on A in all implicit constants (such as the asymptotic notations \ll and $O(\cdot)$) which we will not always explicitly state.

We let Γ denote a convex, symmetric function $\Gamma : \mathbb{R} \rightarrow \mathbb{R}^+$ such that $\Gamma(0) = 0$, Γ is increasing on \mathbb{R}^+ , and $\Gamma(t)$ tends to infinity as t tends to infinity. The (Luxemburg) Orlicz space norm associated to Γ is then defined by:

Definition 2. For $f : \mathbb{T} \rightarrow \mathbb{R}$,

$$\|f\|_{\Gamma_K} := \inf \left\{ \gamma > 0 \mid \int_{\mathbb{T}} \Gamma_K \left(\frac{f}{\gamma} \right) \leq 1 \right\}.$$

We define the following convex function from \mathbb{R} to \mathbb{R} , parameterized by a value $1 < K < \infty$ and a value $2 < p < 3$:

$$\Gamma_K(t) := \begin{cases} |t|^p, & \text{if } |t| \leq K; \\ \left(1 + \frac{p-2}{2}\right) K^{p-2} t^2 - \left(\frac{p-2}{2}\right) K^p, & \text{if } |t| > K. \end{cases}$$

We also define

$$\gamma_K(t) := \begin{cases} |t|^{p-2}, & \text{if } |t| \leq K; \\ K^{p-2}, & \text{if } |t| > K. \end{cases}$$

We observe that $t^2 \gamma_K(t) \leq \Gamma_K(t) \leq \left(1 + \frac{p-2}{2}\right) t^2 \gamma_K(t)$.

Lemma 3. For any fixed $1 < K < \infty$ and $2 < p < 3$, it holds for all $s, t \in \mathbb{R}$ that

$$|s\gamma_K(s) - t\gamma_K(t)| \leq 3(\gamma_K(s) + \gamma_K(t))|s - t|.$$

Proof. Without loss of generality, we may assume that $|s| \geq |t|$. We define β by $t = \beta s$, where $|\beta| \leq 1$. For notational concision, we also define $\alpha = p - 2$. We first consider the case where $|s|, |t| \leq K$. In this case, we consider the quantity

$$|s\gamma_K(s) - t\gamma_K(t)| = |s|^{1+\alpha} (1 - \beta|\beta|^\alpha).$$

We must establish that this is upper bounded by $3|s|^{1+\alpha}(1 - \beta)(1 + |\beta|^\alpha)$. Since $|\beta| \leq 1$ and $\alpha \leq 1$, we have $|\beta|^\alpha \geq |\beta|$, and hence $(1 - \beta)(1 + |\beta|^\alpha) = 1 - \beta + |\beta|^\alpha - \beta|\beta|^\alpha \geq 1 - \beta|\beta|^\alpha$. The desired inequality follows (note that the factor of 3 is not needed for this case).

We next consider the case where $|s| > K$ and $|t| \leq K$. Here, we wish to bound the quantity $|sK^\alpha - t|t|^\alpha|$ by the quantity $3|s - t|(K^\alpha + |t|^\alpha)$. We suppose that $s \geq K \geq t \geq 0$. Then, it suffices to show that

$$\frac{sK^\alpha - t^{1+\alpha}}{s - t} \leq 3K^\alpha,$$

which holds if and only if

$$\frac{s - t^{1+\alpha}K^{-\alpha}}{s - t} \leq 3.$$

We define $0 \leq c \leq 1$ by $t = cK$ and $\sigma > 0$ by $s = (1 + \sigma)K$. Then we have:

$$\frac{s - t^{1+\alpha}K^{-\alpha}}{s - t} = \frac{1 + \sigma - c^{\alpha+1}}{1 + \sigma - c}.$$

We observe:

$$\frac{1 + \sigma - c^{\alpha+1}}{1 + \sigma - c} = \frac{1 - c^{1+\alpha}}{1 + \sigma - c} + \frac{\sigma}{1 + \sigma - c} \leq \frac{1 - c^{1+\alpha}}{1 - c} + \frac{\sigma}{\sigma}.$$

Since $\alpha \leq 1$ and $c \leq 1$, we note that $c^{1+\alpha} \geq c^2$, so this is $\leq 1 + c + 1 \leq 3$, as required.

We now suppose instead that $s \geq K$, and $-K \leq t < 0$. In this case, we need to show that

$$\frac{sK^\alpha + |t|^{1+\alpha}}{s + |t|} \leq 3K^\alpha.$$

Dividing out K^α , we obtain the equivalent requirement

$$\frac{s + |t|^{1+\alpha}K^{-\alpha}}{s + |t|} \leq 3.$$

Using that $|t| \leq K \leq s$, we observe that $\frac{s + |t|^{1+\alpha}K^{-\alpha}}{s + |t|} \leq \frac{2s}{s} = 2$, and so the desired inequality holds. The cases where $s < -K$ and $|t| \leq K$ can be handled symmetrically.

Finally, we must consider the case where $|s|, |t| \geq K$. In this case, we wish to bound the quantity

$$|sK^\alpha - tK^\alpha| = K^\alpha|s - t| = \frac{1}{2}(\gamma_K(s) + \gamma_K(t))|s - t|,$$

so this is clearly $\leq 3|\gamma_K(s) + \gamma_K(t)| \cdot |s - t|$ as required. \square

Lemma 4. For $K \geq 1$, for all t we have that

$$\begin{aligned} \Gamma_K(t) &\leq t^p \\ \Gamma_K(t) &\leq \left(1 + \frac{p-2}{2}\right) K^{p-2}t^2. \end{aligned}$$

It follows that for $f : \mathbb{T} \rightarrow \mathbb{R}$ we have $\|f\|_{\Gamma_K} \leq \|f\|_p$ and $\|f\|_{\Gamma_K} \leq \left(1 + \frac{p-2}{2}\right)^{1/2} K^{(p-2)/2} \|f\|_2$.

Proof. The inequality $\Gamma_K(t) \leq \left(1 + \frac{p-2}{2}\right) K^{p-2}t^2$ is clear from the definition of $\Gamma_K(t)$. We prove $\Gamma_K(t) \leq t^p$. This is clear for $|t| \leq K$, so we assume $|t| > K$. We let $0 \leq c \leq 1$ be defined by $c := \frac{K}{t}$. It then suffices to show

$$\left(1 + \frac{p-2}{2}\right) c^{p-2} - \left(\frac{p-2}{2}\right) c^p \leq 1.$$

Setting the derivative (with respect to c) equal to 0, we see that

$$0 = \frac{d}{dc} \left(1 + \frac{p-2}{2}\right) c^{p-2} - \left(\frac{p-2}{2}\right) c^p = (p-2) \left(1 + \frac{p-2}{2}\right) c^{p-3} - \left(\frac{p-2}{2}\right) p c^{p-1}.$$

This implies that $\left(1 + \frac{p-2}{2}\right) c^{p-3} = \frac{p}{2} c^{p-1}$ and it follows that $c = 1$. Lastly the inequality can be easily verified at $c = 0, 1$.

Let f be a function from \mathbb{T} to \mathbb{R} , and let $\lambda = \|f\|_p$. Now, using $\Gamma_K(t) \leq t^p$,

$$\int \Gamma_K(f/\lambda) \leq \|f\|_p^{-p} \int |f|^p \leq 1.$$

Similarly, setting $\lambda = \left(1 + \frac{p-2}{2}\right)^{1/2} K^{(p-2)/2} \|f\|_2$ and using $\Gamma_K(t) \leq \left(1 + \frac{p-2}{2}\right) K^{p-2}t^2$, we have

$$\int \Gamma_K(f/\lambda) \leq \left(\left(1 + \frac{p-2}{2}\right)^{1/2} K^{(p-2)/2} \|f\|_2 \right)^{-2} \int \left(1 + \frac{p-2}{2}\right) K^{p-2} |f|^2 \leq 1.$$

\square

Given Γ_K , we also define its *dual*, $\Gamma_K^* : \mathbb{R} \rightarrow \mathbb{R}$, as

$$\Gamma_K^*(x) = \int_0^{|x|} (\Gamma_K')^{-1}(t) dt.$$

By a straightforward computation, we have for $t \geq 0$:

$$(\Gamma_K')^{-1}(t) = \begin{cases} (t/p)^{\frac{1}{p-1}}, & \text{if } t \leq pK^{p-1}; \\ t/(pK^{p-2}), & \text{if } t \geq pK^{p-1}. \end{cases}$$

We then compute that for $0 \leq x \leq pK^{p-1}$

$$\Gamma_K^*(x) = \int_0^x \left(\frac{t}{p}\right)^{\frac{1}{p-1}} dt = p^{-\frac{p}{p-1}}(p-1)x^{\frac{p}{p-1}}. \quad (1)$$

For $x > pK^{p-1}$, we compute

$$\Gamma_K^*(x) = \int_0^{pK^{p-1}} \left(\frac{t}{p}\right)^{\frac{1}{p-1}} dt + \int_{pK^{p-1}}^x \frac{t}{pK^{p-2}} dt = \frac{x^2}{2pK^{p-2}} + \left(\frac{p-2}{2}\right) K^p. \quad (2)$$

We call Γ_K^* the dual of Γ_K because $\|\cdot\|_{\Gamma_K^*}$ is the equivalent to the dual norm of $\|\cdot\|_{\Gamma_K}$. More precisely:

Lemma 5. *There exist (universal) positive constants C_1, C_2 such that, for all f ,*

$$C_1 \sup_{\|g\|_{\Gamma_K^*} \leq 1} \int fg \leq \|f\|_{\Gamma_K} \leq C_2 \sup_{\|g\|_{\Gamma_K^*} \leq 1} \int fg.$$

This follows from the standard theory of Orlicz spaces (see Chapter 2 of [7], for instance).

Lemma 6. *For any measurable $f : \mathbb{T} \rightarrow \mathbb{R}$, we can decompose $f = f_1 + f_2$ such that*

$$\|f_1\|_{L^p} \ll \|f\|_{\Gamma_K} \text{ and } \|f_2\|_{L^2} \ll K^{(2-p)/2} \|f\|_{\Gamma_K}.$$

Proof. By homogeneity we may assume that $\|f\|_{\Gamma_K} = 1$. We let \mathbb{I}_S denote the indicator function of a set $S \subset \mathbb{T}$. We now define

$$f_1 := f \cdot \mathbb{I}_{\{x: |\frac{f(x)}{2}| < K\}} \text{ and } f_2 := f \cdot \mathbb{I}_{\{x: |\frac{f(x)}{2}| \geq K\}}.$$

Using the hypothesis that $\|f\|_{\Gamma_K} = 1$, we have

$$\int \Gamma_K(f/2) = \int (f/2)^p \mathbb{I}_{\{x: |\frac{f(x)}{2}| < K\}} + \int \left(\left(1 + \frac{p-2}{2}\right) K^{p-2} (f/2)^2 - \left(\frac{p-2}{2}\right) K^p \right) \mathbb{I}_{\{x: |\frac{f(x)}{2}| \geq K\}} \leq 1.$$

It follows that

$$\int (f/2)^p \mathbb{I}_{\{x: |\frac{f(x)}{2}| < K\}} \leq 1,$$

which is equivalent to $\|f_1\|_p \leq 2$ (or $\|f_1\|_p \leq 2\|f\|_{\Gamma_K}$). Next we have that

$$\begin{aligned} & \int \left(\left(1 + \frac{p-2}{2}\right) K^{p-2} (f/2)^2 - \left(\frac{p-2}{2}\right) K^p \right) \mathbb{I}_{\{x: |\frac{f(x)}{2}| \geq K\}} \\ &= \int \left(1 + \frac{p-2}{2}\right) K^{p-2} (f/2)^2 \mathbb{I}_{\{x: |\frac{f(x)}{2}| \geq K\}} - \mu \left(\left\{ x : \left| \frac{f(x)}{2} \right| \geq K \right\} \right) \frac{p-2}{2} K^p \leq 1, \end{aligned}$$

where μ denotes the Lebesgue measure.

Since we are assuming $\|f\|_{\Gamma_K} = 1$ and $\Gamma_K(f(x)/2) \geq K^p$ whenever $|f(x)/2| \geq K$, we must have $\mu\left(\left\{x : \left|\frac{f(x)}{2}\right| \geq K\right\}\right) \leq K^{-p}$. Combining this with the above inequality, we see that

$$\int \left(1 + \frac{p-2}{2}\right) K^{p-2} (f/2)^2 \mathbb{I}_{\{x: |f(x)/2| \geq K\}} \leq 1 + \frac{p-2}{2}.$$

This implies

$$\int (f_2)^2 \leq 4K^{2-p}.$$

□

Lemma 7. *We have that $\|f\|_2 \ll pK^{p-1}\|f\|_{\Gamma_K^*}$.*

Proof. Referring to the computation of Γ_K^* in (2), we see that

$$\Gamma_K^*(x) \geq \mathbb{I}_{\geq pK^{p-1}}(x) \cdot \frac{x^2}{2pK^{p-2}}.$$

Let $\|f\|_{\Gamma^*} = 1$, then

$$\frac{1}{2pK^{p-2}} \int_{\mathbb{T}} \mathbb{I}_{\geq pK^{p-1}}(f) f^2 \leq \int_{\mathbb{T}} \Gamma_K^*(f) \leq 1. \quad (3)$$

We also note that

$$\|f\|_{L^2}^2 = \int_{\mathbb{T}} \mathbb{I}_{\leq pK^{p-1}}(f) f^2 + \int_{\mathbb{T}} \mathbb{I}_{\geq pK^{p-1}}(f) f^2 \leq p^2 K^{2p-2} + \int_{\mathbb{T}} \mathbb{I}_{\geq pK^{p-1}}(f) f^2. \quad (4)$$

Combining (3) and (4), we have $\|f\|_{L^2}^2 \ll p^2 K^{2p-2}$.

□

We will now recall several probabilistic results that we will need later. The following lemma (see p. 229 of [2]) is the chaining argument from Bourgain's work on the $\Lambda(p)$ -problem:

Lemma 8. *Let $\mathcal{E} \subset \mathbb{R}^N$ and $B := \sup_{x \in \mathcal{E}} \|x\|_{\ell^2}$ (the diameter of \mathcal{E}). Let $0 < \delta < 1$ and $\{\xi_i\}_{i=1}^N$ independent $0, 1$ -valued random variables (selectors) with mean $\delta = \int \xi_i(\omega) d\omega$ and $1 \leq m \leq N$. Then for any $q \geq 1$,*

$$\left\| \sup_{x \in \mathcal{E}, |S| \leq m} \sum_{i \in S} \xi_i(\omega) x_i \right\|_{L^q(d\omega)} \ll \left[\delta m + \frac{q}{\log(1/\delta)} \right]^{1/2} + \log^{-1/2}(1/\delta) \int_0^B \log^{1/2}(N_2(\mathcal{E}, t)) dt. \quad (5)$$

Here, $N_2(\mathcal{E}, t)$ denotes the entropy number of \mathcal{E} with respect to the ℓ^2 -distance. In other words, $N_2(\mathcal{E}, t)$ is the minimum number of ℓ^2 -balls of radius t needed to cover the set \mathcal{E} .

The following lemma (see p. 231 of [2]) is the entropy estimate from Bourgain's work on the $\Lambda(p)$ -problem. There it is stated for orthonormal systems uniformly bounded by 1, but it generalizes easily to those uniformly bounded by a fixed constant A :

Lemma 9. *Let $\Phi := \{\phi_i\}_{i=1}^N$ denote an orthonormal system of functions uniformly bounded by A . Further let $1 \leq m \leq N$ and $2 \leq q < \infty$ and define*

$$\mathcal{P}_m := \left\{ \sum_{i \in S} a_i \phi_i : \sum a_i^2 \leq 1, |S| \leq m \right\}.$$

Then, for some $v_q > 2$,

$$\log(N_q(\mathcal{P}_m, t)) \ll_A C_q m \log\left(\frac{N}{m} + 1\right) t^{-v_q} \text{ for } t > \frac{1}{2} \quad (6)$$

$$\log(N_q(\mathcal{P}_m, t)) \ll_A C_q m \log\left(\frac{N}{m} + 1\right) \log(1/t) \text{ for } 0 < t \leq \frac{1}{2}. \quad (7)$$

The following lemma appears as Theorem 5.2 in [16] (see also [17] and [18]). This deep and general result is a key technical component of Talagrand's alternate approach to the $\Lambda(p)$ -problem and is closely related to what is often called the majorizing measures theorem.

Lemma 10. *Consider an operator U from ℓ_N^2 to the Banach space of real valued functions on \mathbb{T} with a norm $\|\cdot\|$. Further assume that $\|\cdot\| \leq \|\cdot\|_p$ for some $p > 2$. Now consider N independent mean δ selectors $\{\xi_i\}_{i=1}^N$ and consider the random subset $S := \{i \in [N] : \xi_i = 1\}$. Then, if we denote by $U|_S$ the restriction of domain of the operator U to sequences supported on S , we have that*

$$\mathbb{E}[\|U|_S\|^2] \ll \frac{\|U\|^2 + C_p \log(N)}{\log(1/\delta)}. \quad (8)$$

The following lemma appears as Theorem 7.4 in [8] and provides a strong concentration bound for an empirical processes. This is also due to Talagrand [15], although from work on a somewhat different topic.

Lemma 11. *Let Y_1, \dots, Y_N denote random variables taking values in a Banach space W , and \mathcal{F} a countable collection of measurable functions on W pointwise bounded by C (i.e. $|f| \leq C$). Set $Z := \sup_{f \in \mathcal{F}} \sum_{i=1}^N f(Y_i)$ and $\sigma^2 = \sup_{f \in \mathcal{F}} \sum_{i=1}^N |f(Y_i)|^2$. Then, for all $\tau > 0$,*

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq \tau) \leq 3 \exp\left(-\frac{\tau}{\kappa C} \log\left(1 + \frac{C\tau}{\mathbb{E}\sigma^2}\right)\right). \quad (9)$$

for some absolute constant κ .

We will also need the following standard fact:

Lemma 12. *Let $(V_y)_{y \in E}$ be a set of real, non-negative random variables indexed by a finite set E of size N . Let $q \geq 1$. Then*

$$\left\| \sup_y V_y \right\|_q \ll \sup_y \|V_y\|_{q+\log N}.$$

Proof. We define $Z := \sup_y V_y^q$ and let $\eta \geq 1$ be a parameter to be set later. Applying Jensen's inequality, we have $(\mathbb{E}[Z])^\eta \leq \mathbb{E}[Z^\eta]$. We note that $Z^\eta \leq \sum_y V_y^{q\eta}$, so by the linearity of expectation we have

$$(\mathbb{E}[Z])^\eta \leq \sum_y \mathbb{E}[V_y^{q\eta}] \leq N \sup_y \|V_y\|_{q\eta}^{q\eta}.$$

We then observe:

$$\left\| \sup_y V_y \right\|_q = (\mathbb{E}[Z])^{\frac{1}{q}} \leq N^{\frac{1}{q\eta}} \sup_y \|V_y\|_{q\eta}.$$

We then choose η so that $\eta q = q + \log N$. We then have $N^{\frac{1}{q\eta}} \leq N^{\frac{1}{\log N}} \ll 1$, so the lemma follows. \square

Finally, we recall Pittel's inequality, which can be found on p.38 in [1], for example:

Lemma 13. For positive integers $1 \leq m \leq N$, we let $\binom{[N]}{m}$ denote the set of all subsets of $[N]$ of size m . We let S_m denote a uniformly random element of $\binom{[N]}{m}$. We let \tilde{S} denote a subset of $[N]$ chosen by including each element of $[N]$ independently with probability $\delta = m/N$. Then, for any event E described as a set of subsets of $[N]$, we have

$$\mathbb{P} \left[S_m \in E \cap \binom{[N]}{m} \right] \ll \sqrt{m} \cdot \mathbb{P} [\tilde{S} \in E].$$

3 Dyadic Decompositions

Let $\Phi = \{\phi_i\}_{i=1}^N$ be an orthonormal system and fix $f = \sum_{n \in [N]} a_n \phi_n$ in the span of the system. We now perform a dyadic decomposition of f in terms of the ℓ^2 weight of the coefficients (which we refer to as ‘mass’). We reproduce the description of this decomposition from [10].

We define the *mass* of a subinterval $I \subseteq [N]$ as $M(I) := \sum_{n \in I} |a_n|^2$. By normalization, we may assume that $M([N]) = 1$. We define $I_{0,1} := [N]$ and we iteratively define $I_{k,s}$ for $1 \leq s \leq 2^k$ as follows. Assuming we have already defined $I_{k-1,s}$ for all $1 \leq s \leq 2^{k-1}$, we will define $I_{k,2s-1}$ and $I_{k,2s}$, which are subintervals of $I_{k-1,s}$. $I_{k,2s-1}$ begins at the left endpoint of $I_{k-1,s}$ and extends to the right as far as possible while covering strictly less than half the mass of $I_{k-1,s}$, while $I_{k,2s}$ ends at the right endpoint of $I_{k-1,s}$ and extends to the left as far as possible while covering at most half the mass of $I_{k-1,s}$. More formally, we define $I_{k,2s-1}$ as the maximal subinterval of $I_{k-1,s}$ which contains the left endpoint of $I_{k-1,s}$ and satisfies $M(I_{k,2s-1}) < \frac{1}{2}M(I_{k-1,s})$. We also define $I_{k,2s}$ as the maximal subinterval of $I_{k-1,s}$ which contains the right endpoint of $I_{k-1,s}$ and satisfies $M(I_{k,2s}) \leq \frac{1}{2}M(I_{k-1,s})$. We note that these subintervals are disjoint. We may express $I_{k-1,s} = I_{k,2s-1} \cup I_{k,2s} \cup i_{k,s}$, where $i_{k,s} \in I_{k-1,s}$. In other words, $i_{k,s}$ denotes the single element which lies between $I_{k,2s-1}$ and $I_{k,2s}$ (note that such a point always exists because we have required that $I_{k,2s-1}$ contains strictly less than half of the mass of the interval). Here it is acceptable for some choices of the intervals in this decomposition to be empty. By construction we have that

$$M(I_{k,s}) \leq 2^{-k}. \tag{10}$$

For $J \subseteq [N]$, we define

$$f_J(x) = \sum_{n \in J} a_n \phi_n(x).$$

We also define

$$\tilde{f}_J(x) := \max_{I \subseteq J} \left| \sum_{n \in I} a_n \phi_n(x) \right|.$$

where $\max_{I \subseteq J}$ denotes the maximum over subintervals.

Lemma 14. Let $\Phi = \{\phi_i\}_{i=1}^N$ denote an orthonormal system such that $|\phi_i(x)| \leq A$ for all i, x and for every $f = \sum_{n \in [N]} a_n \phi_n$ in the span of the system one has

$$\|f\|_{\Gamma_K} \leq \Delta \|f\|_2$$

for some constant $\Delta \geq 2$. Then we may decompose the maximal function $\tilde{f} = f_1 + f_2$ where

$$\|f_1\|_p \ll_p \Delta \|f\|_2$$

$$\|f_2\|_2 \ll_p \Delta \log(N) K^{(2-p)/2} \|f\|_2.$$

Proof. We normalize $\|f\|_2 = 1$ and split $f = f_L + f_S$ where $f_L := \sum_{|a_n| > N^{-1}} a_n \phi_n$ and $f_S := \sum_{|a_n| \leq N^{-1}} a_n \phi_n$. Now $\tilde{f} \leq \tilde{f}_L + \tilde{f}_S$ and $\tilde{f}_S \leq \sum_{|a_n| \leq N^{-1}} |a_n| |\phi_n| \leq A$. Clearly $\|\tilde{f}_S\|_p \ll \|f\|_2$, so it remains to prove the desired properties for \tilde{f}_L .

We next perform the ‘mass’ decomposition described above on f_L . On each interval $I_{k,s}$ in the dyadic decomposition of f_L , we can apply Lemma 6 and the hypothesis $\|f\|_{\Gamma_K} \leq \Delta \|f\|_2$ to express the restriction of f_L to $I_{k,s}$ as the sum of two functions $G_{k,s}$ and $E_{k,s}$ such that $\|G_{k,s}\|_p \ll \Delta \|f_{I_{k,s}}\|_2$ and $\|E_{k,s}\|_2 \leq \Delta K^{(2-p)/2} \|f_{I_{k,s}}\|_2$. Now, at each point x , the maximizing subinterval achieving the value $\tilde{f}_L(x)$ can be decomposed into a disjoint union of dyadic intervals (and points) using at most two intervals on each level. We then have the pointwise inequality

$$\begin{aligned} \tilde{f}_L &\ll \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |G_{k,s} + E_{k,s}|^p \right)^{1/p} + \sum_k \left(\sum_s |f_{i_{k,s}}|^p \right)^{1/p} \\ &\ll \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |G_{k,s}|^p \right)^{1/p} + \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |E_{k,s}|^p \right)^{1/p} + \sum_k \left(\sum_s |f_{i_{k,s}}|^p \right)^{1/p} \end{aligned}$$

where we have used the condition $|a_n| \geq N^{-1}$ to restrict the sum in k to $\lceil \log(N) \rceil$ values. (We have also used that taking the ℓ_p -norm of the values of the intervals on a single level provides an upper bound on the highest value taken on that level.)

Note that it suffices to prove that an appropriate decomposition exists for a pointwise majorant. We treat each of the three sums on right of the above inequality separately. By hypothesis, we have that $\|G_{k,s}\|_p \leq \Delta \|f_{I_{k,s}}\|_2$ and $\|E_{k,s}\|_2 \leq \Delta K^{(2-p)/2} \|f_{I_{k,s}}\|_2$. We then see that

$$\begin{aligned} \left\| \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |G_{k,s}|^p \right)^{1/p} \right\|_p &\leq \sum_{k=1}^{\lceil \log(N) \rceil} \left\| \left(\sum_s |G_{k,s}|^p \right)^{1/p} \right\|_p \\ &= \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s \|G_{k,s}\|_p^p \right)^{1/p} \leq \sum_{k=1}^{\lceil \log(N) \rceil} \Delta 2^{k(1/p-1/2)} \ll_p \Delta. \end{aligned}$$

Here we have used that $\|f_{I_{k,s}}\|_2 \leq 2^{-k/2}$.

Next we have

$$\begin{aligned} \left\| \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |E_{k,s}|^p \right)^{1/p} \right\|_2 &\leq \left\| \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s |E_{k,s}|^2 \right)^{1/2} \right\|_2 \\ &\leq \sum_{k=1}^{\lceil \log(N) \rceil} \left(\sum_s \|E_{k,s}\|_2^2 \right)^{1/2} \ll \Delta \log(N) K^{(2-p)/2}. \end{aligned}$$

Recall that $f_{i_{k,s}} = a_j \phi_j$ for some $j \in [N]$. By hypothesis we have $|\phi_j(x)| \leq A$ for all x , furthermore $|a_j| \leq 2^{-(k-1)/2}$ since $j \in I_{k-1,s'}$ for some s' . Thus $|f_{i_{k,s}}| \ll_A 2^{-(k-1)/2}$. Hence, for any x ,

$$\sum_k \left(\sum_s |f_{i_{k,s}}(x)|^p \right)^{1/p} \ll \sum_k \left(\sum_s 2^{-(p-2)(k-1)/2} |f_{i_{k,s}}(x)|^2 \right)^{1/p} \ll \sum_k 2^{-(p-2)(k-1)/(2p)} \ll_p 1.$$

Since this term is pointwise bounded, clearly its L^p norm is also bounded. Since the decomposition holds for each of the three terms it clearly holds for their sum, completing the proof. \square

We now record the following easily verified fact (see Lemma 29 in [10]).

Lemma 15. *We fix $\|f\|_2 = 1$ (in the span of $\{\phi_i\}_{i=1}^N$) and let \mathcal{A} denote the set of intervals that occur in the mass decomposition of f . For every interval $J \subseteq [N]$, ($J \neq \emptyset$), there exist $\tilde{J}_\ell, \tilde{J}_r \in \mathcal{A}$ and $i_J \in [N]$ such that $\tilde{J} := \tilde{J}_\ell \cup i_J \cup \tilde{J}_r$ is an interval (i.e. $\tilde{J}_\ell, i_J, \tilde{J}_r$ are adjacent), $J \subseteq \tilde{J}$, and $M(\tilde{J}) \leq 2M(J)$.*

Lemma 16. *Let $\Phi = \{\phi_i\}_{i=1}^N$ denote a bounded orthonormal system. Furthermore, let $M < N$, let $C', r > 0$ be positive constants, and let $K \geq 2$. We assume that for any interval $I \subseteq [N]$ of length $|I| \leq M$ we have for any $h = \sum_{n \in I} a_n \phi_n$, the estimate*

$$\|h\|_{\Gamma_K} \leq \frac{C' \log^r(N)}{\log^r(N/|I|)} \|h\|_2$$

holds. In addition, we assume that

$$\|\mathcal{M}f\|_2 \ll \log \log(N) \|f\|_2$$

holds for any f in the span of the system. Then, if $f = \sum_{n \in \mathcal{I}} a_n \phi_n$ for $|\mathcal{I}| = M \leq N$, for any $\beta > 0$ we have

$$\begin{aligned} \|\mathcal{V}^2 f\|_2 &\ll_p (\log^{\beta/2}(N) + (C')^{p/2} \log^{rp/2 - \beta(p-2)/4}(N) \log^{(1-pr)/2}(N/M) \\ &\quad + C' \log^{r+1}(N) K^{(2-p)/2} + \log \log(N)) \|f\|_2. \end{aligned} \quad (11)$$

Proof. Let $f = \sum_{n \in \mathcal{I}} a_n \phi_n$ with $|\mathcal{I}| = M$ and $\mathcal{I} \subseteq [N]$. Normalize $\|f\|_2 = 1$, and again split $f = f_L + f_S$ where $f_L = \sum_{|a_n| > N^{-1}} a_n \phi_n$ and $f_S = \sum_{|a_n| \leq N^{-1}} a_n \phi_n$. We then have $\mathcal{V}^2 f_S \leq \sum |a_n \phi_n| \ll 1$ and may restrict our attention to f_L .

We now perform a mass decomposition on f_L , using the notation above. Let I be a dyadic subinterval. We consider the restriction of f_L to this interval, and we apply Lemma 14 to decompose the maximal version of this restriction as a sum of functions $\tilde{G}_I + \tilde{E}_I$. We may apply Lemma 14 to the system $\{\phi_i\}_{i \in I}$ with $\Delta := \frac{C' \log^r(N)}{\log^r(N/|I|)}$. We then obtain

$$\|\tilde{G}_I\|_p \ll_p \frac{C' \log^r(N)}{\log^r(N/|I|)} \|f_I\|_2,$$

$$\|\tilde{E}_I\|_2 \ll_p \frac{C' \log^{r+1}(N)}{\log^r(N/|I|)} K^{(2-p)/2} \|f_I\|_2.$$

For a fixed $x \in \mathbb{T}$, we know that the value of $\mathcal{V}^2 f_L(x)$ is achieved by some maximizing partition π . Appealing to Lemma 15, we know that each interval in π can be covered by dyadic intervals and points (and each dyadic interval/point will only be used to cover at most a constant number of the original intervals). Let \mathcal{A} denote the family of dyadic intervals and points. We write $\pi^* \subset \mathcal{A}$ to denote a suitable covering of a partition by dyadic intervals and points. For each k , we define $I_k^a := \{I_{k,s} \text{ s.t. } |I_{k,s}| \leq 2^{-k/2} M\}$ and $I_k^b := \{I_{k,s} \text{ s.t. } |I_{k,s}| > 2^{-k/2} M\}$. We let I^a denote the union of I_k^a over the values of k let I^b denote the union of I_k^b . We then have

$$\begin{aligned} |\mathcal{V}^2 f_L(x)|^2 &\ll \sup_{\pi} \sum_{I \in \pi} |f_I(x)|^2 \ll \sup_{\pi^* \subset \mathcal{A}} \left(\sum_{I \in \pi^*} |\tilde{f}_I|^2 + \sum_{i_{k,s} \in \pi^*} |f_{i_{k,s}}|^2 \right) \\ &\ll \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\tilde{f}_I|^2 + \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^b}} |\tilde{f}_I|^2 + \sup_{\pi^* \subset \mathcal{A}} \sum_{i_{k,s} \in \pi^*} |f_{i_{k,s}}|^2. \end{aligned}$$

The third quantity here can be easily handled. Since $f_{i_{k,s}} = a_{i_{k,s}} \phi_{i_{k,s}}$ and $\sum_i a_i^2 = 1$, the contribution from the points here is $\ll 1$. The second sum may also be estimated in a crude manner. Recall that $\|\tilde{f}_{I_{k,s}}\|_2^2 \ll (\log \log(N))^2 2^{-k}$ (by assumption) and note that $|I_k^b| \leq 2^{k/2}$. Thus

$$\int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^b}} |\tilde{f}_I|^2 \ll \sum_k \sum_{I_{k,s} \in I_k^b} (\log \log(N))^2 2^{-k} \ll (\log \log(N))^2 \sum_k 2^{-k/2} \ll (\log \log(N))^2.$$

We return to the first sum. We fix $\beta > 0$. For each dyadic interval I , we define the ‘bad’ set $B_I = \{x \in \mathbb{T} : |\tilde{G}_I(x)|^2 \geq \log^\beta(N) M(I)\}$ and we let H_I denote its complement inside \mathbb{T} . Then,

$$\begin{aligned} & \int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\tilde{f}_I|^2 \ll \int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\tilde{G}_I + \tilde{E}_I|^2 \\ & \ll \int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\mathbb{I}_{H_I} \tilde{G}_I|^2 + \int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\mathbb{I}_{B_I} \tilde{G}_I|^2 + \int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\tilde{E}_I|^2. \end{aligned}$$

By Definition, $|\mathbb{I}_{H_I} \tilde{G}_I(x)|^2 \leq \log^\beta(N) M(I)$ for all x , and since the intervals in the partition π are disjoint, we have that $\sum_{I \in \pi^*} M(I) \ll 1$. Thus

$$\int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\mathbb{I}_{H_I} \tilde{G}_I(x)|^2 \leq \log^\beta(N).$$

Next, using that $\int |E_{k,s}|^2 \ll_p (C')^2 \log^{2(r+1)}(N) \log^{-2r}(N/|I_{k,s}|) K^{2-p} M(I_{k,s})$, we have

$$\int \sup_{\pi^* \subset \mathcal{A}} \sum_{\substack{I \in \pi^* \\ I \in I^a}} |\tilde{E}_I|^2 \ll_p (C')^2 \log^{2r+2}(N) K^{2-p}.$$

It remains to estimate the contribution of the bad sets. By Chebyshev’s inequality, we have

$$\int \mathbb{I}_{B_{k,s}} \leq \frac{\int |\tilde{G}_{k,s}|^p}{\log^{\beta p/2}(N) (M(I_{k,s}))^{p/2}} \ll_p (C')^p \log^{pr}(N) \log^{-pr}(N/|I_{k,s}|) \log^{-\beta p/2}(N).$$

Now, by Holder’s inequality, we have

$$\int |\mathbb{I}_{B_{k,s}} \tilde{G}_{k,s}|^2 \leq \left(\int \mathbb{I}_{B_{k,s}} \right)^{1/(p/2)'} \left(\int |\tilde{G}_{k,s}|^p \right)^{2/p}$$

where $(p/2)'$ denotes the conjugate exponent of $p/2$. This is

$$\begin{aligned} & \ll_p \left((C')^p \log^{pr}(N) \log^{-pr}(N/|I_{k,s}|) \log^{-\beta p/2}(N) \right)^{1/(p/2)'} (C')^2 \log^{2r}(N) \log^{-2r}(N/|I_{k,s}|) M(I_{k,s}) \\ & = (C')^p \log^{pr}(N) \log^{-pr}(N/|I_{k,s}|) \log^{-\beta(p-2)/2}(N) M(I_{k,s}), \end{aligned}$$

using that $1/(p/2)' = (p-2)/p$. Since $|I_{k,s}| \leq 2^{-k/2} M$ for $I_{k,s} \in I^a$, we have that (assuming $pr > 1$)

$$\begin{aligned} & \int \sum_{k,s, I_{k,s} \in I^a} |\mathbb{I}_{B_{k,s}} \tilde{G}_{k,s}|^2 \ll_p (C')^p \log^{pr}(N) \log^{-\beta(p-2)/2}(N) \left(\sum_{k=1}^{\log(N)} (\log(N/M) + k)^{-pr} \right) \\ & \ll_p (C')^p \log^{pr}(N) \log^{-\beta(p-2)/2}(N) \int_{\log(N/M)}^{\infty} x^{-pr} dx \ll_p (C')^p \log^{pr}(N) \log^{-\beta(p-2)/2}(N) \log^{1-pr}(N/M). \end{aligned}$$

□

Corollary 17. *With the same hypothesis as Lemma 16, if $C' \log^{r+1}(N) K^{(2-p)/2} \ll 1$, $\log(N/M) \ll \log^\theta(N)$, and $r + \theta/p - \theta r > 0$, then*

$$\|\mathcal{V}^2 f\|_2 \ll_{p,C'} \log^{r+\theta/p-\theta r}(N) \|f\|_2$$

for all $f = \sum_{n \in \mathcal{I}} a_n \phi_n$ with $|\mathcal{I}| = M$.

Proof. By hypothesis, the third term in (11) is $\ll 1$, so we only need to choose a $\beta > 0$ to balance the first two terms. Solving $\beta/2 = rp/2 - \beta(p-2)/4 + \theta(1-pr)/2$ we see that $\beta = 2(r + \theta/p - \theta r)$. Taking this choice of β in (11) yields the corollary. \square

At the beginning of this section we introduced a ‘mass’ dyadic decomposition of $[N]$ with respect to a function $f = \sum_{n \in [N]} a_n \phi_n$. Now we recall the more common ‘length’ dyadic decomposition. Without loss of generality, we assume $N = 2^\ell$ for some positive integer ℓ (if N is not a power of 2 we will simply round up to the nearest power of 2). Consider the collection of dyadic subintervals of the form $\mathcal{I}_{k,s} = (s2^k, (s+1)2^k]$ for each $0 \leq k \leq \ell$, $0 \leq s \leq 2^{\ell-k} - 1$. Note that we have used a slightly different indexing convention here, compared with the mass decomposition.

Lemma 18. *Let $J \subseteq [N]$ be an arbitrary subinterval. Then we may decompose $J = J_l \cup J_r$ as a union of disjoint intervals J_l, J_r where (i) at least one of J_l and J_r is non-empty, and (ii) $J_l \subseteq \mathcal{I}_{k,s}$ for some k, s where $|J_l| \geq \frac{1}{2} |\mathcal{I}_{k,s}|$ (assuming J_l is non-empty), and the same holds for J_r .*

Proof. We define $b \in J$ to be of the form $m2^k$ where k is maximal. In other words, we consider all multiples of powers of 2 inside J , and we set b to be one of the multiples of the highest power appearing. We set $J_l := \{1, \dots, b\} \cap J$ and $J_r := \{b+1, \dots, N\} \cap J$. We consider $|J_l|$ and we let k_l be the minimal integer such that $2^{k_l} \geq |J_l|$. Then we must have $k_l \leq k$ (otherwise, a multiple of 2^{k+1} would appear in J , contradicting maximality of k). Thus, there is a dyadic interval of length 2^{k_l} that covers J_l and has length $\leq 2|J_l|$. An analogous argument applies to J_r . \square

As above, if $f(x) = \sum_{n \in [N]} a_n \phi_n(x)$ is fixed and $I \subseteq [N]$ is an interval we will write $f_I = \sum_{n \in I} a_n \phi_n$ and $\tilde{f}_I = \max_{J \subseteq I} |\sum_{n \in J} a_n \phi_n(x)|$ (where the maximum is over all subintervals of I).

Lemma 19. *Let $\Phi = \{\phi_i\}_{i=1}^N$ denote a bounded orthonormal system such that*

$$\|\mathcal{M}f\|_2 \ll \log \log(N) \|f\|_2$$

holds for all f in the span of Φ . Suppose we have $0 < \theta < 1$, $r > 0$, and $p > 2$ such that for each subinterval $\mathcal{I} \subseteq [N]$ of length $M = 2^m$ where M is the largest power of 2 that is $\leq N2^{-\log^\theta(N)}$,

$$\|\mathcal{V}^2 f\|_2 \ll \log^{r+\theta/p-\theta r}(N) \|f\|_2$$

holds for any $f = \sum_{n \in I} a_n \phi_n$. Then, for any $f = \sum_{n \in [N]} a_n \phi_n$, we have that

$$\|\mathcal{V}^2 f\|_2 \ll (\log^{\theta/2}(N) \log \log(N) + \log^{r+\theta/p-\theta r}(N)) \|f\|_2.$$

In particular, when $\theta = \frac{r}{1/2+r-1/p}$, we obtain

$$\|\mathcal{V}^2 f\|_2 \ll \log^{\frac{r}{1+2r-2/p}}(N) \log \log(N) \|f\|_2.$$

Proof. Setting $f(x) = \sum_{n \in [N]} a_n \phi_n(x)$, we may normalize $\|f\|_2^2 = \sum_{n \in [N]} a_n^2 = 1$. We now wish to estimate the quantity

$$\int \max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{n \in I} a_n \phi_n(x) \right|^2 \quad (12)$$

where \mathcal{P}_N ranges over the set of all partitions of $[N]$. From the elementary inequality $(a+b)^2 \leq 3(a^2 + b^2)$ we have that $|\sum_{n \in J} a_n \phi_n(x)|^2 \ll |\sum_{n \in J_l} a_n \phi_n(x)|^2 + |\sum_{n \in J_r} a_n \phi_n(x)|^2$ whenever $J = J_l \cup J_r$ for disjoint intervals J_l, J_r . Using Lemma 18, it follows that we may restrict the set of partitions \mathcal{P}_N in (12) to the subclass \mathcal{P}_N^* of permutations where each interval in each partition is contained in a dyadic interval of at most twice its length. That is,

$$(12) \ll \int \max_{\pi \in \mathcal{P}_N^*} \sum_{I \in \pi} \left| \sum_{n \in I} a_n \phi_n(x) \right|^2.$$

For a fixed x , let $\pi(x)$ denote the partition in \mathcal{P}_N^* achieving the maximum in the above expression at x . We then have that the above quantity can be expressed as

$$\int \sum_{\substack{I \in \pi(x) \\ |I| \geq N2^{-2 \log^\theta(N)}}} \left| \sum_{n \in I} a_n \phi_n(x) \right|^2 + \int \sum_{\substack{I \in \pi(x) \\ |I| < N2^{-2 \log^\theta(N)}}} \left| \sum_{n \in I} a_n \phi_n(x) \right|^2 := B_1 + B_2$$

We now consider the contribution from the quantity B_1 . Let $I \in \pi(x)$, since $\pi(x) \in \mathcal{P}_N^*$ there is a dyadic interval J_I , such that $|I| \leq |J_I| \leq 2|I|$. Since the intervals in $\pi(x)$ are disjoint, and the associated dyadic interval J_I has length at most $2|I|$, it follows that any particular dyadic interval is associated to at most 2 intervals in $\pi(x)$. We recall $\ell = \log(N)$. It follows that

$$B_1 \ll \sum_{0 \leq k \leq 2\ell^\theta} \sum_{0 \leq s < 2^{\ell-k}} |\tilde{f}_{\mathcal{I}_{k,s}}(x)|^2 \ll \log^\theta(N) (\log \log(N))^2, \quad (13)$$

where we have used that $\int |\tilde{f}_I|^2 \ll \int |\mathcal{M}f_I|^2 \ll (\log \log(N))^2 \sum_{n \in I} a_n^2$.

Next, we must estimate the quantity B_2 . Let $M = 2^m$ denote the largest power of 2 less than or equal to $N2^{-\log^\theta(N)}$. Thus $\frac{1}{2}N2^{-\log^\theta(N)} < M \leq N2^{-\log^\theta(N)}$. Consider the partition of N into (dyadic) subintervals of length M , $\{\mathcal{I}_{m,s}\}_{s=0}^{2^{\ell-m}-1}$. It now follows from the hypothesis that on every interval in this partition $\mathcal{I}_{m,s}$, we have $\|\mathcal{V}^2 f\|_2 \ll \log^{r+\theta/p-\theta r}(N) \|f\|_2$ for any $f = \sum_{n \in \mathcal{I}_{m,s}} a_n \phi_n$.

Next we note that each of the intervals in the partitions in B_2 is contained in a dyadic interval of length at most $2 \times N2^{-2 \log^\theta(N)} \leq M$. By the nesting of dyadic intervals, it follows that each interval in the sum defining B_2 is strictly contained in an element of the partition $\{\mathcal{I}_{m,s}\}$. Thus one has

$$B_2 \ll \sum_s |\mathcal{V}^2 f_{\mathcal{I}_{m,s}}|^2 \ll \log^{2(r+\theta/p-\theta r)}(N) \|f\|_2^2. \quad (14)$$

Combining 13 and 14 (and taking square roots) we have that

$$\|\mathcal{V}^2 f\|_2 \ll (\log^{\theta/2}(N) \log \log(N) + \log^{r+\theta/p-\theta r}(N)) \|f\|_2.$$

□

4 Probabilistic Estimates

4.1 A First Estimate

We begin by establishing a close variant of Lemma 3.4 in [4], mainly following the proof in [4] having taking more care with logarithmic factors.

Lemma 20. *Let ϕ_1, \dots, ϕ_N be orthonormal functions on a probability space \mathbb{T}, μ satisfying $|\phi_i| \leq A$ everywhere on \mathbb{T} for all i . Let $0 < \delta < 1$ and let $\{\xi_i\}_{i=1}^N$ be independent, $\{0, 1\}$ -valued random variables of mean δ on a probability space Ω . Let $1 < K < \infty$, $2 < p < 3$, $1 \leq m \leq N$ and $q_0 \geq 1$. Then there exists a*

$$\lambda \ll_p \delta^{\frac{1}{4}} K^{\frac{p-2}{2}} + \left(\frac{q_0}{\log(\frac{1}{\delta})} \right)^{\frac{1}{4}} + \left(\frac{\log N}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2p}}$$

such that

$$\left\| \sup_{\substack{|S| \leq m \\ |c_i|^2 < 2/m}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_{i \in S} c_i \xi_i(\omega) \phi_i}{\lambda} \right) \right\|_{L^{q_0}(d\omega)} \leq 1.$$

Proof. For a fixed $\lambda > 0$ (to be set later) and any fixed $\omega \in \Omega$ and W of size $\leq m$ and coefficients c_i satisfying $|c_i|^2 < 2/m$, we observe that

$$\int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_{i \in W} c_i \xi_i(\omega) \phi_i}{\lambda} \right) \ll \lambda^{-2} \cdot \frac{1}{\sqrt{m}} \sup_{g \in \mathcal{P}_m} \sum_{i \in W} \xi_i(\omega) |\langle \phi_i, g \gamma_K(\lambda^{-1} g) \rangle|, \quad (15)$$

where

$$\mathcal{P}_m := \left\{ g = \sum_{i \in S} a_i \phi_i \mid \sum_i |a_i|^2 \leq 1 \text{ and } |S| \leq m \right\}.$$

We define $\mathcal{E} \subseteq \mathbb{R}^N$ as

$$\mathcal{E} := \left\{ (|\langle \phi_i, g \gamma_K(\lambda^{-1} g) \rangle|)_{i=1}^N \mid g \in \mathcal{P}_m \right\}.$$

We also define $B := \sum_{x \in \mathcal{E}} \|x\|_{\ell^2}$.

Then, applying Lemma 8, we have

$$\begin{aligned} & \left\| \sup_{g \in \mathcal{P}_m, |W| \leq m} \sum_{i \in W} \xi_i(\omega) \cdot |\langle \phi_i, g \gamma_K(\lambda^{-1} g) \rangle| \right\|_{L^{q_0}(d\omega)} \\ & \ll \left(\delta m + \frac{q_0}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2}} B + \left(\log \left(\frac{1}{\delta} \right) \right)^{-\frac{1}{2}} \int_0^B (\log N_2(\mathcal{E}, t))^{\frac{1}{2}} dt, \end{aligned} \quad (16)$$

where $N_2(\mathcal{E}, t)$ denotes the entropy number of \mathcal{E} with respect to the ℓ^2 -distance.

We now derive an upper bound on B for our set \mathcal{E} . We note that for any $g \in \mathcal{P}_m$, we can apply Bessel's inequality to obtain

$$\left(\sum_i |\langle \phi_i, g \gamma_K(\lambda^{-1} g) \rangle|^2 \right)^{\frac{1}{2}} \leq \left(\int_{\mathbb{T}} |g|^2 |\gamma_K(\lambda^{-1} g)|^2 \right)^{\frac{1}{2}}.$$

Using the fact that $|\phi_i| \leq A$, we have that for all $x \in \mathbb{T}$, $|g(x)| \ll \sum_i |a_i| \leq \sqrt{m}$ by the Cauchy-Schwarz inequality. Thus, since $\|g\|_{L^2} \leq 1$, we have $B \ll \gamma_K \left(\frac{\sqrt{m}}{\lambda} \right)$.

Next we address the quantity $N_2(\mathcal{E}, t)$. For arbitrary $g, h \in \mathcal{P}_m$, we consider the quantity

$$\left(\sum_i (|\langle \phi_i, g\gamma_K(\lambda^{-1}g) \rangle| - |\langle \phi_i, h\gamma_K(\lambda^{-1}h) \rangle|)^2 \right)^{\frac{1}{2}} \leq \left(\sum_i |\langle \phi_i, g\gamma_K(\lambda^{-1}g) - h\gamma_K(\lambda^{-1}h) \rangle|^2 \right)^{\frac{1}{2}},$$

using the fact that for any real numbers a and b , $||a| - |b|| \leq |a - b|$. By Bessel's inequality, this quantity is $\leq \|g\gamma_K(\lambda^{-1}g) - h\gamma_K(\lambda^{-1}h)\|_{L^2}$.

Applying Lemma 3, we see this is

$$\ll \| |g - h| \cdot |\gamma_K(\lambda^{-1}g) + \gamma_K(\lambda^{-1}h)| \|_{L^2} \leq \lambda^{-(p-2)} \| |g - h| \cdot (|g|^{p-2} + |h|^{p-2}) \|_{L^2}.$$

Noting that $|g|^{p-2} + |h|^{p-2} \leq 2(|g| + |h|)^{p-2}$, this is $\ll \lambda^{-(p-2)} \| |g - h| \cdot (|g| + |h|)^{p-2} \|_{L^2}$. Now applying Hölder's inequality with conjugate exponents $\frac{1}{3-p}$ and $\frac{1}{p-2}$, we see this quantity is

$$\ll \lambda^{-(p-2)} \left(\int_{\mathbb{T}} |g - h|^{\frac{2}{3-p}} \right)^{\frac{3-p}{2}} \cdot \left(\int_{\mathbb{T}} (|g| + |h|)^2 \right)^{\frac{p-2}{2}} \ll \lambda^{-(p-2)} \|g - h\|_{\frac{2}{3-p}}.$$

By a change of variable, we then have

$$\int_0^\infty (\log N_2(\mathcal{E}, t))^{\frac{1}{2}} dt \ll \lambda^{-(p-2)} \int_0^\infty \left(\log N_{\frac{2}{3-p}}(\mathcal{P}_m, z) \right)^{\frac{1}{2}} dz. \quad (17)$$

Here, $N_{\frac{2}{3-p}}(\mathcal{P}_m, z)$ denote the corresponding entropy numbers of \mathcal{P}_m considered as a subset of the space $L^{\frac{2}{3-p}}(\mathbb{T}, \mu)$.

Applying Lemma 9, we obtain

$$\int_0^\infty \left(\log N_{\frac{2}{3-p}}(\mathcal{P}_m, t) \right)^{\frac{1}{2}} dt \leq C_p \sqrt{m} \left(\log \left(\frac{N}{m} + 1 \right) \right)^{\frac{1}{2}} \left(\int_0^{\frac{1}{2}} \sqrt{\log \left(\frac{1}{t} \right)} + \int_{\frac{1}{2}}^\infty t^{-\frac{\nu_p}{2}} dt \right),$$

where C_p and ν_p denote values that depend only on p . We note that $\nu_p > 2$ for $2 < p < 3$. Thus, we deduce

$$\int_0^\infty \left(\log N_{\frac{2}{3-p}}(\mathcal{P}_m, t) \right)^{\frac{1}{2}} dt \leq C_p \sqrt{m} \left(\log \left(\frac{N}{m} + 1 \right) \right)^{\frac{1}{2}} \quad (18)$$

for some constant C_p depending only on p (we have abused notation a bit here, as this is not the same C_p as in the previous statement).

Combining (17) and (18), we see that

$$\int_0^\infty (\log N_2(\mathcal{E}, t))^{\frac{1}{2}} dt \leq C_p \cdot \lambda^{-(p-2)} \cdot \sqrt{m} \left(\log \left(\frac{N}{m} + 1 \right) \right)^{\frac{1}{2}}. \quad (19)$$

We can now use this to obtain the following bound on the righthand side of (16):

$$\ll \left(\delta m + \frac{q_0}{\log \left(\frac{1}{\delta} \right)} \right)^{\frac{1}{2}} \gamma_K \left(\frac{\sqrt{m}}{\lambda} \right) + C_p \cdot \lambda^{-(p-2)} \cdot \sqrt{m} \left(\log \left(\frac{1}{\delta} \right) \right)^{-\frac{1}{2}} \left(\log \left(\frac{N}{m} + 1 \right) \right)^{\frac{1}{2}}.$$

Combining this with (15), we conclude that

$$\left\| \sup_{\substack{|W| \leq m \\ |c_i| < 2/m}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_{i \in W} c_i \xi_i(\omega) \phi_i}{\lambda} \right) \right\|_{L^{q_0}(d\omega)} \\ \ll \lambda^{-2} \left(\delta + \frac{q_0}{m \log(\frac{1}{\delta})} \right)^{\frac{1}{2}} \gamma_K \left(\frac{\sqrt{m}}{\lambda} \right) + C_p \cdot \lambda^{-p} \left(\frac{\log(\frac{N}{m} + 1)}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2}}.$$

This quantity will be ≤ 1 for a choice of λ that is

$$\lambda \ll_p \delta^{\frac{1}{4}} K^{\frac{p-2}{2}} + \left(\frac{q_0}{\log(\frac{1}{\delta})} \right)^{\frac{1}{4}} + \left(\frac{\log N}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2p}}.$$

□

4.2 Moving to General Coefficients

Lemma 21. *Let ϕ_1, \dots, ϕ_N be orthonormal functions on a probability space \mathbb{T}, μ satisfying $|\phi_i| \leq A$ everywhere on \mathbb{T} for all i . Let $0 < \delta < 1$ and let $\{\xi_i\}_{i=1}^N$ be independent, $\{0, 1\}$ -valued random variables of mean δ on a probability space Ω . Let $1 < K < \infty$, $2 < p < 3$, $1 \leq M \leq N$ and $q_0 \geq 1$. Then there exists a*

$$\lambda \ll_p \delta^{\frac{1}{4}} K^{\frac{p-2}{2}} \sqrt{\log M} + \left(\frac{q_0 + \log \log M}{\log(\frac{1}{\delta})} \right)^{\frac{1}{4}} \sqrt{\log M} + \left(\frac{\log N}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2p}} \sqrt{\log M}$$

such that

$$\left\| \sup_{\substack{||\{a_i\}||_{\ell^2} \leq 1 \\ |support(\{a_i\})| \leq M}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_i a_i \xi_i(\omega) \phi_i}{\lambda} \right) \right\|_{L^{q_0}(d\omega)} \leq 1.$$

Proof. For any fixed values $\{a_i\}$ with support size $\leq M$ such that $\sum_i a_i^2 \leq 1$, we divide them into $\log M$ levels corresponding to powers of 2. More precisely, for each m that is a power of 2 that is ≥ 1 and $< M$, we let S_m denote the set of indices i such that $1/m \leq |a_i|^2 < 2/m$. For $m = 2^{\lceil \log M \rceil}$, we define S_m to be the set of all indices i such that $|a_i|^2 < 2/m$. We then have that $[N]$ is the disjoint union of S_m as $\log m$ ranges from 1 to $\lceil \log M \rceil$. Note that $|S_m| \leq m$ for each m . We also define $d_m = \sum_{i \in S_m} |a_i|^2$ for each m . We note that $\sum_m d_m \leq 1$. For notational convenience, we also define the quantity $D := \sum_m \sqrt{d_m}$.

We then observe (for a fixed $\omega \in \Omega$)

$$\begin{aligned} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_i a_i \xi_i(\omega) \phi_i}{\lambda} \right) &= \int_{\mathbb{T}} \Gamma_K \left(\lambda^{-1} \sum_m \sum_{i \in S_m} a_i \xi_i(\omega) \phi_i \right) \\ &= \int_{\mathbb{T}} \Gamma_K \left(\sum_m \frac{\sqrt{d_m}}{D} \left(\frac{D}{\lambda \sqrt{d_m}} \sum_{i \in S_m} a_i \xi_i(\omega) \phi_i \right) \right). \end{aligned}$$

Appealing to the convexity of Γ_K and the linearity of the integral, we see this is

$$\leq D^{-1} \sum_m \sqrt{d_m} \int_{\mathbb{T}} \Gamma_K \left(\frac{D}{\lambda \sqrt{d_m}} \sum_{i \in S_m} a_i \xi_i(\omega) \phi_i \right).$$

For each m , we define the random variable $V_{D,m}$ as:

$$V_{D,m}(\omega) = \sup_{\substack{|A| \leq m, \\ |c_i|^2 \leq 2/m}} \int_{\mathbb{T}} \Gamma_K \left(\frac{D}{\lambda} \sum_{i \in A} c_i \xi_i(\omega) \phi_i \right).$$

We now have

$$\begin{aligned} & \left\| \sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ |\text{support}(\{a_i\})| \leq M}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_i a_i \xi_i(\omega) \phi_i}{\lambda} \right) \right\|_{L^{q_0}(d\omega)} \\ & \leq \left\| \sup_{\|\{d_m\|_{\ell^2} \leq 1} D^{-1} \sum_m \sqrt{d_m} V_{D,m} \right\|_{L^{q_0}(d\omega)}. \end{aligned}$$

We note that $D = \sum_m \sqrt{d_m} \ll \sqrt{\log M}$ by the Cauchy-Schwarz inequality (recall there are only $\lceil \log M \rceil$ values of m). Therefore, since $V_{D,m}$ is non-decreasing as a function of D , the quantity above is

$$\ll \left\| \sup_{\|\{d_m\|_{\ell^2} \leq 1} D^{-1} \left(\sum_m \sqrt{d_m} \right) \sup_m V_{D,m} \right\|_{L^{q_0}(d\omega)} \leq \left\| \sup_m V_{\sqrt{\log M}, m} \right\|_{L^{q_0}(d\omega)}.$$

Applying Lemma 12, we see that this is

$$\ll \sup_m \left\| \sup_{\substack{|A| \leq m, \\ |c_i|^2 \leq 2/m}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sqrt{\log M}}{\lambda} \sum_{i \in A} c_i \xi_i(\omega) \phi_i \right) \right\|_{L^{q_0 + \log \log M}(d\omega)}.$$

By setting

$$\lambda \ll_p \delta^{\frac{1}{4}} K^{\frac{p-2}{2}} \sqrt{\log M} + \left(\frac{q_0 + \log \log M}{\log(\frac{1}{\delta})} \right)^{\frac{1}{4}} \sqrt{\log M} + \left(\frac{\log N}{\log(\frac{1}{\delta})} \right)^{\frac{1}{2p}} \sqrt{\log M}$$

and applying Lemma 20, we obtain the result. \square

4.3 Obtaining a Good Partition

Lemma 22. *Let ϕ_1, \dots, ϕ_N be orthonormal functions on a probability space \mathbb{T}, μ satisfying $|\phi_i| \leq A$ everywhere on \mathbb{T} for all i . Let $1 \leq M \leq L \leq N$ and $\gamma > 1$. It holds with probability at least $1 - cN^{-\gamma}$ (for some universal constant c) that a subset I of N of size L chosen uniformly randomly satisfies*

$$\sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ |\text{support}(\{a_i\})| \leq M}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_{i \in I} a_i \phi_i}{\lambda} \right) \leq 2 \quad (20)$$

for a choice of λ that is

$$\lambda \ll_{p,\gamma} \left(\frac{L}{N} \right)^{\frac{1}{4}} K^{\frac{p-2}{2}} \sqrt{\log M} + \left(\frac{\log N}{\log(\frac{N}{L})} \right)^{\frac{1}{4}} \sqrt{\log M} + \left(\frac{\log N}{\log(\frac{N}{L})} \right)^{\frac{1}{2p}} \sqrt{\log M}.$$

Proof. We fix a value of λ suitable to apply Lemma 21 (with $\delta = L/N$ and $q_0 = 2\gamma \log(N)$ fixed). For a real value $t > 1$, we let E_t denote the event that a set S of size of L selected uniformly at random from $[N]$ satisfies

$$\sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ \text{support}(\{a_i\}) \leq M}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_{i \in S} a_i \phi_i}{\lambda} \right) \leq t.$$

Similarly, for independent selectors $\{\xi_i\}$ with mean $\delta = L/N$, we let E'_t denote the event that

$$\sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ \text{support}(\{a_i\}) \leq M}} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_i a_i \xi_i(\omega) \phi_i}{\lambda} \right) \leq t.$$

We begin by obtaining a lower bound on the probability of the event E'_t using Lemma 21. We have by Chebyshev's inequality:

$$\mathbb{P} \left[\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \int_{\mathbb{T}} \Gamma_K \left(\frac{\sum_i a_i \xi_i(\omega) \phi_i}{\lambda} \right) > t \right] \leq t^{-q_0}.$$

Hence, $\mathbb{P}[E'_t] \geq 1 - t^{-q_0}$.

Next we observe the relationship between $\mathbb{P}[E_t]$ and $\mathbb{P}[E'_t]$. By applying Lemma 13 to the complements of the events E_t and E'_t , we have $1 - \mathbb{P}[E_t] \ll \sqrt{L} \cdot (1 - \mathbb{P}[E'_t])$. This implies that

$$\mathbb{P}[E_t] \geq 1 - C\sqrt{L} \cdot t^{-q_0} \geq 1 - C\sqrt{N} \cdot t^{-2\gamma \log(N)}$$

for some positive constant C .

Thus for a uniformly random subset of size L , the probability of the event E_2 , that is that (20) holds, is at least $1 - cN^{-\gamma}$. □

Corollary 23. *Let $\{\phi_i\}_{i=1}^N$ be a bounded orthonormal system. Let $1 \leq L \leq N$ and $\gamma > 1$. It holds with probability at least $1 - cN^{-\gamma}$ (for some universal constant c) that a subset of $S \subseteq [N]$ of size L chosen uniformly randomly satisfies:*

$$\sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ \text{support}(\{a_i\}) \leq M}} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_{(N/L)^{1/(2p-4)}}} \ll_{p,\gamma} \left(\frac{\log N}{\log(\frac{N}{L})} \right)^{\frac{1}{4}} \sqrt{\log M}$$

for all M in the range $1 \leq M \leq L$.

Proof. We apply the previous lemma with parameters $\gamma + 1$ and $K := (\frac{N}{L})^{\frac{1}{2p-4}}$ and employ a union bound over the $\leq N$ values of M . We note that once we have a λ such that

$$\int \Gamma_K \left(\frac{\sum_{i \in S} a_i \phi_i}{\lambda} \right) \leq 2,$$

we can multiply λ by a constant to obtain an upper bound on the Γ_K -norm. □

Corollary 24. *Let $\{\phi_i\}_{i=1}^N$ be a bounded orthonormal system. For any $\gamma > 1$, it holds with probability at least $1 - cN^{-\gamma}$ (for some universal c) that a uniformly randomly selected permutation $\sigma : [N] \rightarrow [N]$ satisfies the following. If I is an subinterval in $[N]$ then (for all $M \leq |I|$) we have*

$$\sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ \text{support}(\{a_i\}) \leq M}} \left\| \sum_{i \in I} a_n \phi_{\sigma(n)} \right\|_{\Gamma_{(N/|I|)^{1/(2p-4)}}} \ll_{p,\gamma} \left(\frac{\log N}{\log(\frac{N}{|I|})} \right)^{\frac{1}{4}} \sqrt{\log M}$$

(uniformly in the choice of interval I and M).

Proof. We apply the previous corollary with the parameter $\gamma+2$ for each interval I with $L := |I|$. We then employ the union bound over the $\sim N^2$ intervals. Note that for each I , the image of I under a randomly chosen permutation is equivalent to a randomly chosen subset of size $|I|$ inside $[N]$. \square

5 Bourgain's Theorem

The estimates in the previous section are not strong enough to deduce Theorem 1, however they do allow us to reprove Bourgain's theorem and show that it holds with large probability, a fact we will require later.

Proposition 25. *Let $\{\phi_n\}_{n=1}^N$ denote a bounded orthonormal system. For any $\gamma > 1$ it holds with probability at least $1 - cN^{-\gamma}$ (for some universal c) that a uniformly randomly selected permutation $\sigma : [N] \rightarrow [N]$ satisfies*

$$\|\mathcal{M}_\sigma f\|_2 \ll_\gamma \log \log(N) \|f\|_2$$

where $\mathcal{M}_\sigma f(x) := \max_{\ell \leq N} \left| \sum_{n=1}^\ell a_n \phi_{\sigma(n)}(x) \right|$ for all $f = \sum_{n \in [N]} a_n \phi_{\sigma(n)}$ in the span of the system.

Proof. We select a permutation σ satisfying the conclusion of Corollary 24. We fix $f = \sum_{n \in [N]} a_n \phi_{\sigma(n)}$ and we consider intervals $I_{k,s}$ and points $i_{k,s}$ in a corresponding mass decomposition of $[N]$. For each k , we define I_k^a to be the collection of intervals $I_{k,s}$ such that $|I_{k,s}| \leq 2^{-k/2}N$ and I_k^b to be the collection of intervals $I_{k,s}$ such that $|I_{k,s}| > 2^{-k/2}N$. We observe that $|I_k^b| \leq 2^{k/2}$.

At each fixed point $x \in \mathbb{T}$, the value of $\mathcal{M}_\sigma f(x)$ is achieved on some subinterval of $[N]$ that can be decomposed into a union of dyadic intervals $I_{k,s}$ and points $i_{k,s}$ such that there is at most one interval $I_{k,s}$ for each value of k . We let $k^* := \lceil 15 \log \log(N) \rceil$. We then have the pointwise inequality

$$\begin{aligned} \mathcal{M}_\sigma f(x) &\ll \sum_k \left(\sum_{\substack{s \\ I_{k,s} \in I_k^b}} |f_{I_{k,s}}|^2 \right)^{1/2} + \sum_{k \leq k^*} \left(\sum_{\substack{s \\ I_{k,s} \in I_k^a}} |f_{I_{k,s}}|^2 \right)^{1/2} \\ &\quad + \sum_k \max_s |f_{i_{k,s}}| + \left(\sum_{\substack{1 \leq s \leq 2^{k^*} \\ I_{k^*,s} \in I_{k^*}^a}} |\tilde{f}_{I_{k^*,s}}|^p \right)^{1/p}. \end{aligned}$$

To see this, note that $\left(\sum_{I_{k,s} \in I_k^b} |f_{I_{k,s}}|^2 \right)^{1/2}$ is an upper bound on the largest value of $|f_{I_{k,s}}|$ over all s such that $I_{k,s} \in I_k^b$ for each k (for example), and the final term above captures the contribution from dyadic subintervals for values of $k > k^*$.

We now consider the L^2 norm of each of these terms. Recall that $\int |f_{I_{k,s}}|^2 \leq 2^{-k}$ and $|I_k^b| \leq 2^{k/2}$, thus we have

$$\left\| \sum_k \left(\sum_{\substack{s \\ I_{k,s} \in I_k^b}} |f_{I_{k,s}}|^2 \right)^{1/2} \right\|_2 \leq \sum_k \left(\sum_{\substack{s \\ I_{k,s} \in I_k^b}} \|f_{I_{k,s}}\|_2^2 \right)^{1/2} \leq \sum_k 2^{-k/4} \ll 1.$$

Next,

$$\left\| \sum_{k \leq k^*} \left(\sum_{\substack{s \\ I_{k,s} \in I_k^a}} |f_{I_{k,s}}|^2 \right)^{1/2} \right\|_2 \leq \sum_{k \leq k^*} \left\| \left(\sum_{1 \leq s \leq 2^k} |f_{I_{k,s}}|^2 \right)^{1/2} \right\|_2 \ll \log \log(N).$$

We also have that, using that $|\phi_n(x)| \leq A$ and $\|f_{i_{k,s}}\|_2^2 \leq 2^{-k+1}$, we have the pointwise bound

$$\sum_k \max_s |f_{i_{k,s}}| \ll \sum_k 2^{-k/2} \ll 1.$$

Finally, we must consider the quantity

$$\left\| \left(\sum_{\substack{1 \leq s \leq 2^{k^*} \\ I_{k^*,s} \in I_{k^*}^a}} |\tilde{f}_{I_{k^*,s}}|^p \right)^{1/p} \right\|_2.$$

Using Corollary 24 and Lemma 14, we may decompose $\tilde{f}_{I_{k^*,s}} = G_{I_{k^*,s}} + E_{I_{k^*,s}}$ where $\|G_{I_{k^*,s}}\|_p \ll_{p,\gamma} \log^{3/4}(N) \|f_{I_{k^*,s}}\|_2$ and $\|E_{I_{k^*,s}}\|_2 \ll_{p,\gamma} K^{(2-p)/2} \log^{7/4}(N) \|f_{I_{k^*,s}}\|_2$ where $K = \left(\frac{N}{|I_{k^*,s}|} \right)^{\frac{1}{2p-4}}$. For $I_{k^*,s} \in I_{k^*}^a$, we have $|I_{k^*,s}| \leq 2^{-k^*/2} N$, so $K \geq 2^{\frac{k^*}{4(p-2)}}$.

Setting $p = 5/2$, we now have

$$\begin{aligned} \left\| \left(\sum_{1 \leq s \leq 2^{k^*}} |G_{I_{k^*,s}}|^p \right)^{1/p} \right\|_2 &\leq \left\| \left(\sum_{1 \leq s \leq 2^{k^*}} |G_{I_{k^*,s}}|^p \right)^{1/p} \right\|_p = \left(\sum_{1 \leq s \leq 2^{k^*}} \|G_{I_{k^*,s}}\|_{5/2}^{5/2} \right)^{2/5} \\ &\ll \log^{3/4}(N) \left(\sum_{1 \leq s \leq 2^{k^*}} \|f_{I_{k^*,s}}\|_2^{5/2} \right)^{2/5}. \end{aligned}$$

Now $\sum_{1 \leq s \leq 2^{k^*}} \|f_{I_{k^*,s}}\|_2^{5/2} \ll 2^{k^*} (2^{-k^*/2})^{5/2} \ll \log^{-15/4}(N)$, which implies the quantity is

$$\ll \log^{3/4}(N) \log^{-3/2}(N) \ll 1.$$

We last consider

$$\begin{aligned} \left\| \left(\sum_{\substack{1 \leq s \leq 2^{k^*} \\ I_{k^*,s} \in I_{k^*}^a}} |E_{I_{k^*,s}}|^p \right)^{1/p} \right\|_2 &\leq \left(\int \sum_{\substack{1 \leq s \leq 2^{k^*} \\ I_{k^*,s} \in I_{k^*}^a}} |E_{I_{k^*,s}}|^2 \right)^{\frac{1}{2}} \\ &\ll \left(\sum_{\substack{1 \leq s \leq 2^{k^*} \\ I_{k^*,s} \in I_{k^*}^a}} \log^{7/2}(N) K^{2-p} \|f_{I_{k^*,s}}\|_2^2 \right)^{\frac{1}{2}}. \end{aligned}$$

We recall that $K = \left(\frac{N}{|I_{k^*,s}|} \right)^{\frac{1}{2(p-2)}}$, so for $I_{k^*,s} \in I_{k^*}^a$, we have $K \geq 2^{\frac{k^*}{4(p-2)}}$. Thus, $K^{\frac{2-p}{2}} \leq 2^{-\frac{k^*}{8}}$, so the quantity above is $\ll \log^{7/4}(N) \log^{-15/8}(N) \ll 1$. This completes the proof. \square

6 Improving the Bound by Passing to Random Subsets

6.1 Bounding the Expectation

For a fixed interval $I \subseteq [N]$ and a bounded ONS $\{\phi_i\}_{i=1}^N$, we define the operator U_I from ℓ_N^2 to the Banach space of real valued functions on \mathbb{T} with norm $\|\cdot\|_{\Gamma_K}$ by mapping a sequence $\{a_i\}_{i=1}^N$ to the function $\sum_{i \in I} a_i \phi_i$.

Lemma 26. *Let $\{\phi_i\}_{i=1}^N$ be a bounded ONS. Let I be a fixed interval satisfying*

$$\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in I} a_i \phi_i \right\|_{\Gamma_K} \leq \lambda \quad (21)$$

for a fixed λ and K . Let $\{\xi_i\}_{i \in I}$ be independent selectors with mean ν . We let $S \subseteq I$ denote the set of indices $i \in I$ such that $\xi_i = 1$. Then

$$\mathbb{E} \left[\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}^2 \right] \ll \frac{\lambda^2 + C_p \log N}{\log\left(\frac{1}{\nu}\right)}.$$

Proof. We apply Lemma 10 to the operator U_I . This yields

$$\mathbb{E} \left[\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}^2 \right] \ll \left(\log\left(\frac{1}{\nu}\right) \right)^{-1} \left(C_p \log N + \sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in I} a_i \phi_i \right\|_{\Gamma_K}^2 \right).$$

Using the interval I satisfies (21), we see this quantity is $\ll \frac{\lambda^2 + C_p \log N}{\log\left(\frac{1}{\nu}\right)}$. □

6.2 Bounding the Concentration

We now consider the concentration of the random variable $\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}$ around its expectation.

Lemma 27. *Let $\{\phi_i\}_{i=1}^N$ be a bounded ONS. Fix $0 < \beta < 1$ and let $L \leq N$, $\delta := L/N$. Then for any $\gamma > 1$ it holds with probability at least $1 - cN^{-\gamma}$ (for some universal c) that a uniformly randomly selected subset S of $[N]$ of size L satisfies*

$$\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K} \ll_{p,\gamma} \frac{\log^{3/4}(N)}{\log^{3/4}\left(\frac{N}{L}\right)} \quad (22)$$

for $K := \min \left\{ \left(\frac{1}{\delta}\right)^{1/(4p-8)}, \log^{\frac{7}{2p-4}}(N) \right\}$ and $2 < p < 3$.

Proof. We will pass to a set of size L in two stages. We first define an intermediate value $L_1 \geq L$ such that $\frac{N}{L_1} = \frac{L_1}{L}$. We will set $K := \min \left\{ \left(\frac{N}{L_1}\right)^{1/(2p-2)}, \log^{\frac{7}{2p-4}}(N) \right\}$. We then apply Corollary 23 to conclude that with sufficiently high probability, say at least $1 - cN^{-(\gamma+2)}$, a random subset W of size L_1 inside $[N]$ will satisfy

$$\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in W} a_i \phi_i \right\|_{\Gamma_K} \ll_{p,\gamma} \left(\frac{\log N}{\log\left(\frac{N}{L_1}\right)} \right)^{\frac{1}{4}} \cdot \sqrt{\log L_1}. \quad (23)$$

We now condition on the above result, and consider choosing a random subset $S \subseteq W$ of size L . We first note that for any such set S ,

$$\left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K} = \sup_{\|g\|_{\Gamma^*} \leq 1} \left\langle \sum_{i \in S} a_i \phi_i, g \right\rangle,$$

where $\|\cdot\|_{\Gamma^*}$ denotes the dual norm to $\|\cdot\|_{\Gamma_K}$. Using the Cauchy-Schwarz inequality, we then see that

$$\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K} = \sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} a_i \langle \phi_i, g \rangle \leq \sup_{\|g\|_{\Gamma^*} \leq 1} \left(\sum_{i \in S} |\langle \phi_i, g \rangle|^2 \right)^{\frac{1}{2}}. \quad (24)$$

In fact, noting that one can set $a_i = \frac{\langle \phi_i, g \rangle}{\sqrt{\sum_{i \in S} \langle \phi_i, g \rangle^2}}$, we see that the inequality in (24) is an equality.

We let C be a threshold parameter that we will specify later. We define the functions $\chi_1, \chi_2 : \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\chi_1(x) = \begin{cases} x, & \text{if } |x| \leq C; \\ 0, & \text{otherwise} \end{cases}, \quad \chi_2(x) = \begin{cases} 0, & \text{if } |x| \leq C; \\ x, & \text{otherwise} \end{cases}$$

We then have

$$\sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} |\langle \phi_i, g \rangle|^2 \leq \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_1(|\langle \phi_i, g \rangle|^2) + \sup_{\|h\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_2(|\langle \phi_i, h \rangle|^2). \quad (25)$$

We will deal separately with the two quantities in (25). To address the first quantity, we will employ Lemma 11. In this case, our random variables Y_1, \dots, Y_N are defined as follows. We let Ω denote the probability space of the independent selectors $\{\xi_i\}_{i \in I}$, each with mean $\nu := \frac{L}{L_1}$. Then each $\omega \in \Omega$ is associated to a subset $S \subseteq I$. (This distribution of S differs of course from selecting a set of size exactly L , but we will analyze the relevant quantity in this case first and then derive a bound for the case of fixed size.) We define $Y_i(\omega)$ to be equal to ϕ_i when $i \in S$, and equal to 0 otherwise (more formally, the constant zero function from \mathbb{T} to \mathbb{C}).

For notational convenience, we defined the random variable $Z : \Omega \rightarrow \mathbb{R}$ by

$$Z(\omega) = \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_1(|\langle \phi_i, g \rangle|^2),$$

where S is determined from ω as described above. Applying Lemma 11, it follows that (for every positive real number τ)

$$\mathbb{P}[|Z - \mathbb{E}[Z]| \geq \tau] \leq 3 \exp \left(\frac{-\tau}{\kappa C} \log \left(1 + \frac{C\tau}{\mathbb{E}[\sigma^2]} \right) \right), \quad (26)$$

where κ is a positive constant, and $\sigma^2 = \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \left| \chi_1(|\langle \phi_i, g \rangle|^2) \right|^2$.

Recalling our definition of Z , we observe

$$Z(\omega) \leq \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} |\langle \phi_i, g \rangle|^2 = \sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}^2.$$

(We have used here that (24) is actually an equality.) Consequently,

$$\mathbb{E}[Z] \leq \mathbb{E} \left[\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}^2 \right].$$

Applying Lemma 26, we have

$$\mathbb{E}[Z] \ll \frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})},$$

where with sufficient probability we have $\lambda \ll_{p,\gamma} \log^{\frac{1}{4}}(N) \log^{-\frac{1}{4}}\left(\frac{N}{L_1}\right) \sqrt{\log L_1}$ from (23). Thus, there exist a constant A_1 such that

$$\mathbb{P} \left[Z \geq \tau + A_1 \cdot \frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})} \right] \leq 3 \exp \left(\frac{-\tau}{\kappa C} \log \left(1 + \frac{C\tau}{\mathbb{E}[\sigma^2]} \right) \right). \quad (27)$$

It remains to bound the quantity

$$\mathbb{E}[\sigma^2] = \mathbb{E} \left[\sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \left| \chi_1 \left(|\langle \phi_i, g \rangle|^2 \right) \right|^2 \right].$$

Employing line (7.10) on p. 141 of [8], we see this is

$$\ll \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in I} \mathbb{E} \left[\left(\chi_1 \left(|\langle \xi_i \phi_i, g \rangle|^2 \right) \right)^2 \right] + C \mathbb{E} \left[\sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_1 \left(|\langle \phi_i, g \rangle|^2 \right) \right]. \quad (28)$$

By removing the cutoff at C , the latter quantity in (28) is

$$\begin{aligned} &\leq C \mathbb{E} \left[\sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} |\langle \phi_i, g \rangle|^2 \right] = C \mathbb{E} \left[\sup_{\|\{a_i\}\|_{\ell^2} \leq 1} \left\| \sum_{i \in S} a_i \phi_i \right\|_{\Gamma_K}^2 \right] \\ &\ll C \left(\frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})} \right) \end{aligned}$$

where we have recalled that (24) is an equality and have applied Lemma 26.

We now consider the first quantity in (28). Using the definition of χ_1 and Lemma 7, we see it is

$$\ll C \sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in I} \mathbb{E} [|\langle \xi_i \phi_i, g \rangle|^2] \ll C\nu \sup_{\|g\|_{\Gamma^*} \leq 1} \|g\|_2^2 \ll C\nu p^2 K^{2p-2}.$$

Putting this together we have, for universal constants A_1 and A_2 , that

$$\begin{aligned} &\mathbb{P} \left[Z \geq \tau + A_1 \cdot \frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})} \right] \\ &\leq 3 \exp \left(\frac{-\tau}{\kappa C} \log \left(1 + \frac{C\tau}{A_2 C\nu p^2 K^{2p-2} + A_2 C \left(\frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})} \right)} \right) \right). \quad (29) \end{aligned}$$

We now set $\tau = A_1 \cdot \frac{\lambda^2 + C_p \log N}{\log(\frac{1}{\nu})}$. Since $K^{2p-2} \leq \frac{1}{\nu}$, we see that the quantity inside the logarithm in (29) above is at least $1 + \alpha\tau$, where α is a positive constant depending on p, γ . Since $\nu \geq \frac{1}{N}$,

it is clear that τ is bounded away from 0. Thus, it suffices to take $C = A_{\gamma,p} \log^{-1}(N)$ (for some constant $A_{\gamma,p}$ dependent only on p and γ).

Using Lemma 13, we can derive an analogous upper bound on the probability of the event that

$$\sup_{\|g\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_1 \left(|\langle \phi_i, g \rangle|^2 \right) \geq 2\tau$$

for a randomly chosen set S of size L (inside W). The extra factor of $\leq \sqrt{N}$ can be easily accommodated by choosing C to be a sufficiently high power of $\log N$.

We note that

$$\tau = A_1 \cdot \frac{\lambda^2 + C_p \log N}{\log \left(\frac{1}{\nu} \right)} \ll_{p,\gamma} \frac{\log^{3/2}(N)}{\log^{3/2} \left(\frac{1}{\nu} \right)} + \frac{\log N}{\log \left(\frac{1}{\nu} \right)}$$

whenever (23) holds. Thus, the square root of this quantity is $\ll_{p,\gamma} \frac{\log^{3/4}(N)}{\log^{3/4} \left(\frac{N}{L} \right)}$ as required.

We return to consider the second quantity in (25). By Bessel's inequality, the number of nonzero terms appearing in the sum is at most $C^{-1} \sup_{\|h\|_{\Gamma^*} \leq 1} \|h\|_{L^2}^2$. From Lemma 7 we see that this is $C^{-1} p^2 K^{2p-2}$.

We then have

$$\sup_{\|h\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_2 \left(|\langle \phi_i, h \rangle|^2 \right) \leq \sup_{\|h\|_{\Gamma^*} \leq 1} \sup_{\substack{I \subseteq W \\ |I| \leq C^{-1} p^2 K^{2p-2}}} \sum_{i \in I} |\langle \phi_i, h \rangle|^2. \quad (30)$$

We next observe that

$$\sup_{\|h\|_{\Gamma^*} \leq 1} \sup_{\substack{I \subseteq W \\ |I| \leq C^{-1} p^2 K^{2p-2}}} \sum_{i \in I} |\langle \phi_i, h \rangle|^2 \leq \sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ |\text{Support}(\{a_i\})| \leq C^{-1} p^2 K^{2p-2}}} \sup_{\|h\|_{\Gamma^*} \leq 1} \left(\sum_{i \in W} a_i \langle \phi_i, h \rangle \right)^2. \quad (31)$$

To see this, note that one can set $a_i := \frac{\langle \phi_i, h \rangle}{\left(\sum_{j \in I} |\langle \phi_j, h \rangle|^2 \right)^{\frac{1}{2}}}$ for all i in a set I of size at most $C^{-1} p^2 K^{2p-2}$.

Combining (30) and (31), we see that

$$\sup_{\|h\|_{\Gamma^*} \leq 1} \sum_{i \in S} \chi_2 \left(|\langle \phi_i, h \rangle|^2 \right) \leq \sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ |\text{Support}(\{a_i\})| \leq C^{-1} p^2 K^{2p-2}}} \sup_{\|h\|_{\Gamma^*} \leq 1} \left(\sum_{i \in W} a_i \langle \phi_i, h \rangle \right)^2.$$

Employing Lemma 5, this is

$$\ll \sup_{\substack{\|\{a_i\}\|_{\ell^2} \leq 1 \\ |\text{Support}(\{a_i\})| \leq C^{-1} p^2 K^{2p-2}}} \left\| \sum_{i \in W} a_i \phi_i \right\|_{\Gamma_K}^2.$$

By (our earlier application of) Corollary 23, we see this is

$$\begin{aligned} & \ll_{p,\gamma} \left(\frac{\log N}{\log \left(\frac{N}{L_1} \right)} \right)^{\frac{1}{2}} \log(C^{-1} p^2 K^{2p-2}) = \left(\frac{\log N}{\log \left(\frac{N}{L_1} \right)} \right)^{\frac{1}{2}} \log(A_{\gamma,p}^{-1} \log(N) p^2 K^{2p-2}) \\ & \ll_{p,\gamma} \left(\frac{\log N}{\log \left(\frac{N}{L_1} \right)} \right)^{\frac{1}{2}} \log \log(N). \end{aligned} \quad (32)$$

Clearly this term is acceptable and completes the proof. \square

6.3 Deriving the Main Theorem

Using Lemma 27 with Corollary 17 allows us to take $r = 3/4$ and $p = 3 - \epsilon'$ in Lemma 19 for any $0 < \epsilon' < 1$. More precisely, we can set $\theta := \frac{3/4}{1/2+3/4-1/(3-\epsilon')} = \frac{3}{5-4/(3-\epsilon')}$. To obtain the hypotheses needed to apply Lemma 19, we employ Corollary 17. To obtain the hypotheses needed to apply this corollary, we employ Lemma 27 with $\delta := 2^{-2 \log^\theta(N)}$. Note here that K will be set so that $\log^{r+1}(N)K^{2-p/2} \ll 1$ as required. We also employ Proposition 25. Then, with ϵ defined by $\frac{9}{22} + \epsilon = \frac{3}{10-8/(3-\epsilon')}$, we obtain:

Theorem 1. *Let $\Phi := \{\phi_n\}_{n=1}^N$ denote a bounded ONS. Let $\epsilon > 0, \gamma > 1$. Then, with probability at least $1 - cN^{-\gamma}$ (for some universal c), for a uniformly random permutation $\sigma : [N] \rightarrow [N]$, the system $\{\phi_\sigma(n)\}_{n=1}^N$ will satisfy*

$$\|\mathcal{V}^2 f\|_2 \ll_{\epsilon, \gamma} \log^{\frac{9}{22} + \epsilon}(N) \|f\|_2.$$

7 Concluding Remarks

We remarked earlier that when proving the main probabilistic estimate, Lemma 27, we were unable to work exclusively in either Bourgain or Talagrand's frameworks, and had to use a hybrid of the two. We briefly expound on these issues and how our approach addresses them.

Bourgain's approach to the $\Lambda(p)$ -problem makes very careful use of the structure of the L^p norms, in particular relying on delicate pointwise inequalities involving the function $x \rightarrow |x|^p$. We have been unable to find appropriate analogs of these arguments in the case of the more awkward Γ_K norms of relevance here. Many of these issues can be avoided by first pigeonholing the coefficients to a single level set (indeed even Bourgain's paper [2] can be substantially simplified in this case, see for instance the argument sketched in [3]) however this introduces a logarithmic factor that is fatal for the current application (although, this is essentially the idea behind the estimates in Section 4.1, and the prior work in the maximal and \mathcal{V}^r cases).

In contrast, Talagrand's methods can be applied to a very general class of norms, including those considered here (as we have partly done via Lemma 10). However, when used in the framework of [16] (such as Theorem 1.2 there), one 'loses' a factor of $N^{-\epsilon}$ in the density of the resulting subsets. In the case of L^p norms and polynomially sparse sets (of relevance in the $\Lambda(p)$ -problem), Talagrand is able to decompose the norm into two parts. On one part, better estimates are true for denser sets (and thus the loss of a $N^{-\epsilon}$ factor is acceptable) and his very general theorem can be applied. The other part can be handled with alternate (and more elementary) methods. Again, we have been unable to find an analog of these arguments in the setting of Γ_K norms and on the denser sets of relevance here.

Our approach has been to use Bourgain's arguments to prove that the Γ_K norm on a random subset of the relevant density is within a logarithmic factor of what is needed (we note that the arguments from section 3 of [16] are too crude for the purposes here). We then use the orthonormal system associated to this set as the starting point for the application of Talagrand's methods, which are invoked by passing to a slightly sparser random subset.

This approach gives satisfactory control over the expectation of the Γ_K norm on the resulting subset, however our application requires more information about the concentration around the expectation. Lemma 11 provides a useful estimate in this respect, however this alone does not seem sufficient for our application. Fortunately, this estimate is insufficient only when some of the coefficients are large and we are able to adequately control the contribution from terms with large coefficients by returning to Bourgain's argument with pigeonholing, since the number of relevant level sets is then reduced. Likely, a better understanding/treatment of these issues will lead to an improved estimate.

Some final remarks:

Remark 28. *As with many orthonormal system results (for instance those of [2] and [4]), our arguments can be easily modified to treat the more general case of Hilbertian systems.*

Remark 29. *A quantitative form of Kolmogorov's rearrangement theorem due to Nakata [12] states that the first N exponentials can be reordered such that there exists an f in their span such that $\|\mathcal{M}f\|_{L^2} \gg \log^{1/4}(N)\|f\|_{L^2}$. From modulation and dilation symmetries, this holds for any set of N exponentials associated to an arithmetic progression. Notice that a random reordering of $[N]$ will contain an increasing arithmetic progression of length at least $\gg \log^c(N)$ (for some absolute constant $c > 0$) in this 'bad' ordering. Thus while a random reordering decreases the norm of the \mathcal{V}^2 operator, it increases the norm of the \mathcal{M} operator to (at least) $(\log \log(N))^{1/4}$, with large probability.*

Remark 30. *In a related paper [11], we have shown that given an arbitrary ONS of length N , one may find an alternate ONS that spans the same space such that $\|\mathcal{V}^2 f\|_{L^2} \ll \sqrt{\log \log(N)}\|f\|_{L^2}$, which is sharp. Where the results in our current work rely on estimates for what may be called selector processes, there the problem can be reduced to estimates for Gaussian processes. Generally, estimates in that setting are stronger and better understood.*

References

- [1] B. Bollobas, Random Graphs, 2nd. Ed., Cambridge University Press (2001).
- [2] J. Bourgain, Bounded orthogonal systems and the $\Lambda(p)$ -set problem. Acta Math. 162 (1989), no. 3–4, 227-245.
- [3] J. Bourgain, $\Lambda(p)$ -sets in analysis: results, problems and related aspects. Handbook of the geometry of Banach spaces, Vol. I, 195232, North-Holland, Amsterdam, 2001.
- [4] J. Bourgain, On Kolmogorov's rearrangement problem for orthogonal systems and Garsia's conjecture. Geometric aspects of functional analysis (1987–88), Lecture Notes in Math., 1376, Springer, Berlin, (1989) 209-250.
- [5] A. Garsia, Existence of almost everywhere convergent rearrangements for Fourier series of L^2 functions. Ann. of Math. (2) 79 (1964) 623-629
- [6] A. Garsia, Topics in almost everywhere convergence. Lectures in Advanced Mathematics, 4 Markham Publishing Co., Chicago, Ill. (1970)
- [7] M. Krasnoselskii, J. Rutickii, Convex functions and Orlicz spaces. Translated from the first Russian edition by Leo F. Boron. P. Noordhoff Ltd., Groningen (1961).
- [8] M. Ledoux, The concentration of measure phenomenon. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, (2001)
- [9] A. Lewko and M. Lewko, An exact asymptotic for the square variation of partial sum processes, arXiv:1106.0783
- [10] A. Lewko and M. Lewko, Estimates for the square variation of partial sums of Fourier series and their rearrangements. J. Funct. Anal. 262 (2012), no. 6, 2561-2607
- [11] A. Lewko and M. Lewko, Orthonormal systems in linear spans, arXiv:1205.2420

- [12] S. Nakata, On the divergence of rearranged Fourier series of square integrable functions, *Acta Sci. Math.* 32 (1971), 59–70.
- [13] R. Oberlin, A. Seeger, T. Tao, C. Thiele, J. Wright, A variation norm Carleson theorem. *J. Eur. Math. Soc.* 14 (2012), no. 2, 421-464
- [14] A. Olevskii, Divergent series for complete systems in L^2 . *Dokl. Akad. Nauk SSSR* 138 (1961) 545-548
- [15] M. Talagrand, New concentration inequalities in product spaces. *Invent. Math.* 126 (1996), no. 3, 505-563
- [16] M. Talagrand, Sections of smooth convex bodies via majorizing measures, *Acta Math.* 175 (1995), no. 2, 273-300
- [17] M. Talagrand, Selecting a proportion of characters, *Israel J. Math.* 108 (1998), 173-191
- [18] M. Talagrand, The generic chaining. Upper and lower bounds of stochastic processes. Springer Monographs in Mathematics. Springer-Verlag, Berlin, (2005)
- [19] S. Taylor, Exact asymptotic estimates of Brownian path variation. *Duke Math. J.* 39 (1972), 219-241

A. Lewko, Microsoft Research
allew@microsoft.com

M. Lewko, Department of Mathematics, University of California, Los Angeles
mlewko@math.ucla.edu